УДК 004.8

Ю.В. Крак, В.А. Кузнецов, А.С. Тернов
Институт кибернетики им. В.М. Глушкова НАН Украины, Украина
Украина, 03680, г. Киев, пр-т. Ак. Глушкова, 40

АНАЛИЗ МИМИЧЕСКИХ ПРОЯВЛЕНИЙ НА ЛИЦЕ ЧЕЛОВЕКА

Iu.V. Krak, V.A. Kuznetsov, A.S. Ternov
V.M. Glushkov Institute of Cybernetics, Ukraine
Ukraine, 03680, Kyiv, Glushkov ave., 40

ANALYSIS OF FACIAL EXPRESSIONS ON A HUMAN FACE

Ю.В. Крак, В.А. Кузнєцов, А.С. Тернов
Інститут кібернетики імені В.М. Глушкова НАН України, Україна
Україна, 03680, г. Київ, пр-т. Ак. Глушкова, 40

АНАЛІЗ МІМІЧНИХ ПРОЯВІВ НА ОБЛИЧЧІ ЛЮДИНИ

В работе описаны шаги наполнения библиотеки видеофрагментов мимических выражений. С помощью алгоритмов компьютерного зрения в потоке видео были получены числовые данные, описывающие изменение состояния мимических проявлений во времени. Для анализа полученных данных было предложено использовать два метода – однослойный перцептрон и перцептрон в комбинации с интегральными преобразованиями: Карунена-Лоэва, сингулярным разложением и преобразованием Фурье. Проведен анализ работы каждого алгоритма на тестовых выборках, содержащих от 3 до 16 классов мимических выражений. Предложены решения, которые могут как повысить качество анализируемых данных, так и улучшить работу алгоритмов идентификации мимических выражений.

Ключевые слова: мимика, алгоритмы классификации, перцептрон, преобразование Карунена-Лоэва

Facial expressions were recorded on video, stored in facial expressions videolibrary and processed by means of computer vision. Numeric values were analyzed by two algorithms – one layer perceptron and perceptron combined with Karhunen-Loeve transform, singular value decomposition, discrete cosine transform and Fourier transform as a source of teach data. Each algorithm was tested on test data sets. Some solutions were proposed in order to improve quality of analyzed data and effectiveness of facial expression classification algorithms.

Key words: facial expressions, classification algorithms, perceptron, Karhunen-Loeve transform

У роботі розглянуто послідовність наповнення бібліотеки відеофрагментів мімічних проявів. За допомогою алгоритмів комп’ютерного зору в потоці відео було отримано числові дані, що описують зміну стану мімічних проявів у часі. Для аналізу отриманих даних було запропоновано застосувати 2 методи – одношаровий перцептрон і перцептрон у поєднанні з інтегральними перетвореннями, такими як перетворення Карунена-Лоева, сингулярний розклад і перетворення Фур’є. Проведено аналіз роботи алгоритмів на тестових даних, що містили від 3 до 16 класів мімічних проявів. Запропоновано рішення, що можуть підвищити якість даних, що аналізуються, а також покращити роботу алгоритмів ідентифікації мімічних проявів.

Ключові слова: міміка, алгоритми класифікації, перцептрон, перетворення Карунена-Лоева

Введение

Мимика — один из каналов передачи информации в разговорной речи и жестовом языке. В связи с интенсивным развитием средств компьютерной обработки данных особый интерес в изучении мимики представляет задача классификации мимических выражений по базовому набору мимических выражений. Особенно она актуальна для задачи автоматического транскрибирования мими-

ческих проявлений на видеофрагментах, содержащих жестовую речь. Мимика изучается исследователями жестового языка наравне с мануальными проявлениями, поскольку содержит грамматические элементы жестовой речи, которые передаются в разговорной речи с помощью интонации и других элементов [1-3].

Ранее задача получения данных была решена в исследовании [4] и [5]. Для этого использовались оптические маркеры, положение которых на лице отслеживается с помощью алгоритмов компьютерного зрения. Наличие инструментария для получения данных позволило приблизить решение задачи классификации, однако сама задача поиска пространства признаков и разделения мимических проявлений, даже на основе априорного знания о разделения мимических выражений на классы, является нетривиальной.

Исследователи подходят к задаче классификации, применяя различные методы, которые основаны на выделении особых областей на лице (глаза, брови, нос, рот) с помощью фильтров Габора [5], получении коэффициентов разложения дискретного косинусного преобразования [6] и характеристических векторов для каждого из изображений, затем полученные вектора анализируются с помощью нейросетей. При использовании однослойных нейросетей (однослойного перцептрона) точность распознавания на контрольной выборке составляла 85% [7]. При увеличении количества слоев нейронной сети достигалась точность от 83,5% до 90% [5,6]. С применением гибридной многослойной эволюционной нейросети (HENN) достигалась точность от 92,3% [6]. В некоторых исследованиях анализировались mocap-данные, полученные с помощью системы захвата движения на основе масок [8, 9]. В первом исследовании [8], на многослойной нейросети достигался результат от 76% (для 3 классов мимических выражений) до 81% (7 классов); во втором [9] от 63,9% на однослойной и от 66,3% на многослойной.

Постановка задачи

Необходимо разработать алгоритм для решения задачи классификации мимических выражений на ограниченном наборе базовых мимических выражений.

Для того чтобы получить и апробировать такой алгоритм, в первую очередь, необходимо решить следующие задачи:

- определить, какой набор мимических проявлений и микровыражений возможно выделить, используя инструментарий получения числовых данных;
- определить, какие данные можно использовать для задачи идентификации мимических проявлений;
- выбрать преобразования, которые позволяют переходить от числовых данных в пространство, в котором эти данные могут быть разделимы;
- построить обучающую и тестовую выборки, полученные на основе числовых данных о мимических проявлениях.

Основные сведения о мимических проявлениях

Под мимическим проявлением подразумевается некоторая мимическая активность на лице, связанная с внутренними переживаниями индивида или выражающая некоторое его отношение к какой-то ситуации [10].

При исследовании лица выделялось большое разнообразие мимических выражений, которые пытались упорядочить определенным образом, например,

построив некую иерархию формирования таких проявлений [10-13]. Отмечается, что большая часть эмоциональных состояний человека и сопровождающих их мимических проявлений имеют некую базовую эмоцию и сопровождающее ее мимическое выражение, таким образом, все проявления являются производными от базовых.

Основным различием между мимическим проявлением, микровыражением и смесью микровыражений является одновременность отдельных компонент в проявлениях и наличие связи мышечной активности с изменением нервной активности. Например, при проявлении состояния эмоции радости индивид будет демонстрировать высокую мозговую активность, а при проявлении состояния усталости, напротив, мозговая активность будет низкой [11].

Мимические проявления отличаются по темпу и степени проявления. Темп проявления как, например, в мимическом выражении «испуг» и «удивление» или «улыбка» и «смех» может быть различным. Различается и сила проявления.

Наличие таких особенностей в мимических проявлениях усложняет задачу идентификации состояния лица и, тем более, разделения слабых и сильных проявлений, как визуально, так и с помощью алгоритмов классификации. В данном исследовании будет рассматриваться ограниченный набор мимических выражений, содержащий представителей базовых мимических выражений с разной степенью проявления.

Библиотека мимических проявлений

В работе [4] была выработана методология наполнения базы данных (БД) видеофрагментов мимических проявлений в жестовой речи и ее структура. Набор мимических выражений, содержащийся в данной БД, был сопоставлен с набором наиболее часто встречающихся мимических выражений в разговорной речи [14]. На основе сопоставления представителей из мимических выражений, содержащихся в разговорной и жестовой речи, был получен сокращенный набор из 58 мимических выражений. Затем, полученный набор, на основе базовых мимических выражений, которые рассматривались в исследованиях [10-14], был разделен на 16 классов: «боль», «широкая улыбка», «шок», «презрение», «отвращение», «гнев», «грусть», «радость», «интерес», «надежда», «обеспокоенность», «озадаченность», «сосредоточенность», «недовольство», «скука», «удивление».

Для каждого из представителей были сформированы текстовые описания, описывающие ситуацию, при которой возникает данное мимическое выражение и вид изменений областей лица, которые при этом наблюдаются, а также указан схематический пример лица человека [14], с соответствующим мимическим проявлением.

На основе полученных текстовых описаний был сформирован сценарий записи видеофрагментов. Данный сценарий предусматривал запись для нескольких актеров последовательности из 2-4 мимических выражений, по 2 примера каждого мимического выражения от состояния покоя до максимума проявления и до состояния покоя. Перед записью актеру показывались примеры мимических выражений, которые будут записываться, и области лица, которые формируют данное мимическое выражение. После чего актер показывал указанную последовательность. Всего было получено 42 сценарных видеозаписи, содержащие 332 примера мимических выражений, записанных для 3 разных актеров. Затем, полученные примеры были верифицированы экспертом. Так из 332 примеров мимических выражений было верифицировано 187 примеров,

которые и сформировали библиотеку мимических проявлений. Кадры из 12 видео с мимическими выражениями, вошедшими в библиотеку видеофрагментов, представлены на рис 1.

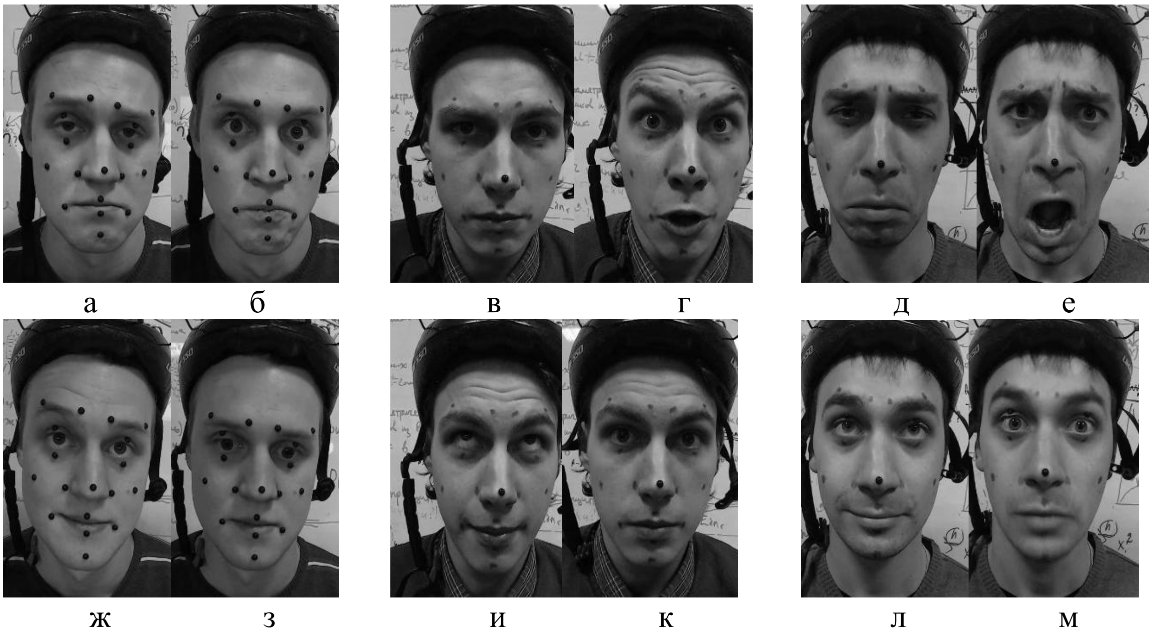


Рис. 1. Мимические выражения: (а, в, д) «грусти», (б, г, е) «обеспокоенности», (ж, и, л) «надежды» и (з, к, м) «удивления» у разных актеров.

Получение числовых данных в видеопотоке

Общие сведения о методах получения числовых данных. Методы получения числовых данных в видеопотоке можно условно разделить на две группы:

-методы получения данных о признаках на основе видеопотока (feature extraction);

-методы получения данных на основе информации о трехмерной структуре исследуемого объекта (3d-mesh capture).

В работе [15] были рассмотрены методы из первой группы, включающие в себя методы идентификации, отслеживания признака и реконструкции положения этого признака в пространстве относительно других признаков.

Для получения данных используются вспомогательные элементы для выделения признаков – оптические датчики (маркеры) [16]. Особенностью таких датчиков является то, что они располагаются в определенных ключевых точках лица, в которых наблюдается максимум интенсивности мимических проявлений [5]. Кроме того, с помощью параметрических моделей [15] можно получить переход между сложной моделью, содержащей трехмерную структуру (набор состояний трехмерной сетки в разные моменты времени) к более простой, содержащей набор ключевых точек (набор маркеров). По сравнению со второй группой методов первая позволяет работать с меньшим объемом данных (используя свойства параметрической модели) и при меньшей размерности данных (2 координаты вместо 3).

Основой технологии является: камера или набор видеокамер (обычно – 2-3 стационарных или одна, закрепленная на голове актера), с помощью которой

получаются видеофрагменты, и алгоритмы компьютерного зрения для получения траектории движения оптических датчиков.

Выбор минимального набора маркеров, используемых для получения числовых данных

В работе [5] было установлено, что маркеры в системе захвата мимических проявлений с использованием технологии захвата движения (mocap) располагаются в областях лица, отвечающих за изменение тех или иных мышц, а также элементов движения, используемых в системе описания мимических проявлений FACS [17].

Используя свойства мышечной структуры и систем координат параметрической модели лица, полученные в работе [5], можно утверждать о том, что какой-либо из признаков можно представить в пространстве локальных координат каждого из признаков параметрической (анимационной) модели.

Для анализа мимических проявлений предложено использовать набор маркеров, полученный в работе [5] (рис. 2).



Рис. 2. Расположение маркеров на лице актера

Получение числовых данных

В работе [5] подробно рассмотрены шаги получения числовых данных с использованием алгоритмов компьютерного зрения. По сравнению с предложенным ранее алгоритмом, были внесены изменения, которые включают двухпроходный алгоритм устранения вибрации камеры. Эти два шага расположены как на этапе после записи (получения видеофрагментов), так и после обработки видеозаписей алгоритмами компьютерного зрения.

Два этапа двухпроходного алгоритма устранения колебания видеокамеры включают в себя:

-обработку изображения с помощью алгоритма цифровой стабилизации [18]);

-обработку полученных данных фильтром Калмана [19] со сглаживанием всплесков длительностью менее 200 мс (величина получена экспериментально).

Указанные шаги реализованы в виде модулей на языке C++ в среде Microsoft Visual Studio с использованием библиотеки OpenCV [20] и в среде PyCharm на языке Python соответственно.

Кроме изменений, касающихся программного обеспечения, которое использовалось в экспериментах, также было изменено аппаратное обеспечение, использовавшееся для записи видео.

Было предложено снимать видео крупным планом, с закреплением видеокамеры на специальном шлеме (рис 3).



Рис. 3. Расположение видеокамеры относительно плоскости лица

Для записи использовалась видеокамера с постоянным фокусным расстоянием Hewlett-Packard Action Camera 150, с разрешением FullHD (1920x1080), с аппаратным сжатием в компрессор H.264.

Камера закреплялась неподвижно относительно головы и на одинаковых расстояниях относительно плоскости лица для всех актеров, которые участвовали в записи видео. Расстояние от камеры до плоскости лица было подобрано экспериментальным путем; при таком расстоянии исчезали искажения (сферическая аберрация), связанные с широким углом камеры.

Видео, полученное с камеры, преобразовывалось в формат Motion JPEG с помощью программного обеспечения ffmpeg/avconv [21], однако, в дополнение к этому, видео дополнительно обрезалось и разрешение уменьшалось до 360x540 (это сделано для того, чтобы ускорить обработку полученного видео алгоритмами компьютерного зрения).

Во время обработки данных была обнаружена возможность получения числовых данных, не используя оптические маркеры. В области рта, в области бровей и некоторых других частей лица с ярко выраженными контурами стало возможным получать числовые данные, используя в качестве ключевых точек элементы лица, формирующие геометрические признаки типа «углы» и «концы отрезков». Очевидно, что сказалось как высокое качество исходных данных (по

сравнению с предыдущими исследованиями), так и использование более новой версии гибридного алгоритма SAD-KLT в программном обеспечении Blender [22], применявшегося ранее.

Анализ мимических проявлений

Данные об объекте. Рассмотрим визуализацию графиков изменения одного из мимических проявлений между состояниями покоя с переходом через состояние насыщения (рис 4). Каждой из кривой соответствует изменение одной из координат оптического маркера относительно состояния покоя. На рис 4 приведен образец сигнала, на котором отмечены изменения координат X маркеров № 4,5,6,8,9,10,12 и координаты Y маркеров № 9,10,11,12,17 и 18.

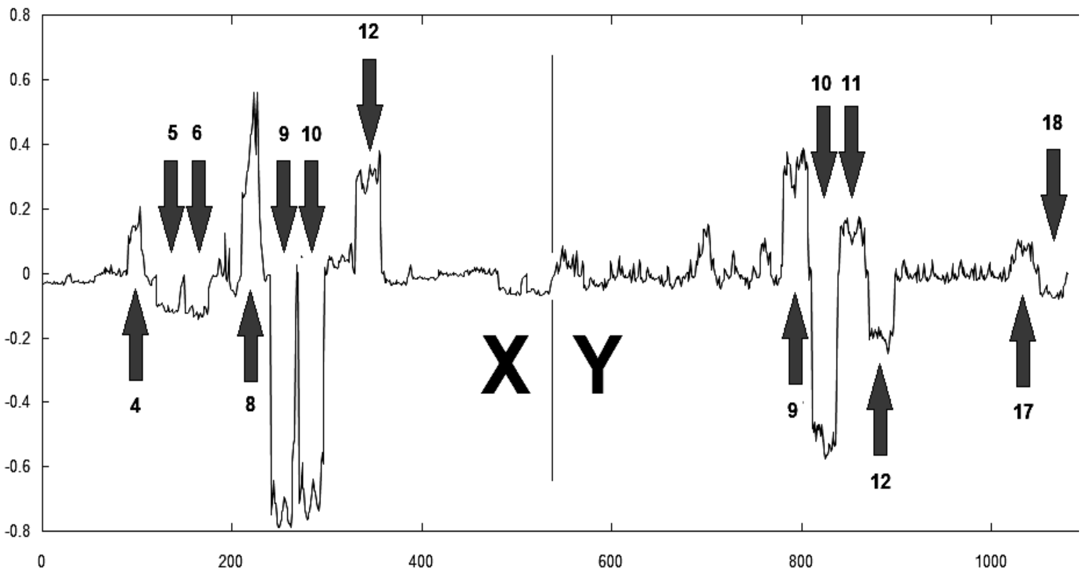


Рис. 4. Расположение активности разных маркеров в сигнале

Как видно из рис. 4, графики представляют собой непериодические импульсные сигналы. Каждый из импульсов в примере графика движения оптического маркера, имеет различные амплитуды и длительности. Кроме того, все видеофрагменты и файлы данных имеют различное количество отсчетов (кадров).

Для задачи анализа мимических проявлений, необходимо определить, что есть элемент данных. Элементом данных может выступать как совокупность всех данных, содержащих траектории движения маркеров (рассматриваем один файл с числовыми данными как объект) или отдельно взятые координаты отдельных маркеров из полученных данных (например, группу маркеров), так и полученный путем усреднения группы элементов, а также часть импульса (фронт, спад, фрагмент сигнала в состоянии насыщения, мгновенное значение в какой-то момент времени).

Зная, что следует выделять в сигнале, как полезную информацию, можно выбрать метод, который наиболее подходит для решения задачи с поставленными условиями.

Выбор метода идентификации мимических проявлений

Для задачи идентификации предложено применить два различных метода.

В первом случае, когда анализируется элемент данных типа «единичный отсчет» было предложено применить однослойный перцептрон.

Во втором случае было предложено использовать различные интегральные преобразования (быстрое преобразование Фурье (FFT) [23], сингулярное разложение (SVD) [24], преобразование Карунена-Лоэва (KLT) [25] и дискретное косинусное преобразование (DCT) [26]) в связке с однослойным перцептроном, который использовался для построения разделяющей гиперплоскости в пространстве признаков каждого из преобразований. Для проверки работы методов результаты сравнивались с результатами работы алгоритма перцептрона, когда сигнал целиком использовался для обучения, без промежуточного преобразования.

Подготовка данных для однослойного перцептрона

Для анализа разных по величине проявления сигналов у разных актеров, было предложено ввести параметр – относительную величину отклонения маркера, полученную на основе величины абсолютного проявления для каждой из координат оптических маркеров для каждого из актеров и отклонений положения маркера от состояния покоя. Для каждого из маркеров и для каждого из актеров введена универсальная шкала, с градацией проявлений от -1 до 1.

Данный шаг необходим, так как для каждого из актеров величина отклонений каждого из маркеров в абсолютном выражении различная, что может повлиять на значение функции активации для мимических проявлений с малой амплитудой и возможности сравнения данных одинакового масштаба.

После получения данных в таком формате, были сформированы обучающие и контрольная выборки. Для этого для всего набора из 187 мимических проявлений были получены интервалы активности мимических проявлений, которые включали в себя только те отсчеты, в которых все маркеры пребывали в состоянии насыщения.

На основе интервалов активности данные разбивались в соотношении 1 к 5 (5 частей на обучение, 1 часть на проверку), причем отбирались не все данные, а только каждый третий элемент данных (отсчет). Полученная обучающая и проверочная выборки формировались в виде двух отдельных текстовых файлов, которые затем зачитывались алгоритмом во время этапа обучения и проверки.

Было отдельно сформировано 2 разных набора обучающих и тестовых примеров для случая 2 базовых мимических выражений (размер выборки составил около 800 примеров) и для 5 базовых выражений (размер выборки составил около 1700 примеров). Такое различие размеров выборок объясняется тем, что во второй набор вошли примеры мимических выражений, представителей которых не было в 1 наборе.

Подготовка данных для интегральных преобразований

В случае применения интегральных преобразований, окно должно быть фиксированной формы, а сигнал отцентрирован относительно окна. Это необходимое требование, т.к. сигналы с разным коэффициентом заполнения, разной фазой и разной длительностью будут иметь отличающиеся коэффициенты интегральных преобразований.

Данные, полученные на этапе «подготовка данных для однослойного перцептрона», приводились к одинаковой длине путем сжатия или растяжения

сигнала каждой из координат каждого маркера до длины 12,30,60 и 120 отсчетов. Вычисление значений для сигнала новой длины выполнялось с кубического сплайна, построенного на исходных сигналах. Также были сформированы сигналы, полученные путем отсечения от сигнала первых 60 отсчетов и сигналы на основе единичных отсчетов.

Реализация инструментария анализа полученных числовых данных

Для реализации инструментария анализа были разработаны два модуля в среде GNU Octave [27]. Первый модуль реализует обучение и тестирование однослойного перцептрона на базе логистической регрессии. Второй модуль реализует формирование собственных векторов преобразования KLT на основе анализируемого сигнала, на основе коэффициентов разложения SVD каждого из сигналов и на основе коэффициентов разложения сигнала с помощью FFT или DCT, а также собственных векторов преобразования SVD на основе анализируемого сигнала; визуализацию данных в пространстве признаков данного преобразования.

В первом модуле, по количеству уникальных вхождений в файл, содержащий столбец значений функции, вычислялось количество классов в обучающей выборке и выполнялось обучение логистической регрессии в реализации 1 против всех (в данном случае 1 против 1 и 1 против 4). Полученные весовые коэффициенты использовались для проверки алгоритма обучения на основе логистической регрессии; в качестве результата использовалась величина процента правильно определенных файлов из обучающей и тестовой выборок. При наличии примеров, которые неправильно определялись алгоритмом, возвращались номера (массив номеров) файлов для обучающей и тестовой выборки.

Во втором программном модуле файл, содержащий обучающую выборку (в зависимости от данных, это мог быть как сам сигнал, так и отсчеты сигнала и коэффициенты одного из интегральных преобразований), записывался в матрицу прямоугольного вида, где вектор-столбец соответствовал одному примеру из обучающей выборки. Значение длины вектора-столбца было кратно 36 и количеству используемых отсчетов. Для 12, 30, 60 и 120 она составляла 432, 1080, 2160, 4320 (для исходного сигнала и/или коэффициентов разложения FFT/DCT), 1296 (коэффициенты разложения SVD) элементов.

В случае вычисления преобразования KLT, для всей обучающей выборки, вычислялось среднее значение. Полученное среднее вычиталась из всех векторов-столбцов, соответствующих обучающим примерам.

На основе полученной матрицы вычислялись коэффициенты интегрального преобразования Карунена-Лоэва. Для этого вычислялись собственные вектора ковариационной матрицы. Коэффициенты преобразования Карунена-Лоэва получались как произведение ковариационной матрицы, деленной на квадратный корень по модулю диагональной матрицы собственных чисел ковариационной матрицы.

В случае вычисления преобразования SVD, для всей обучающей выборки, вычислялись три матрицы разложения SVD – матрица собственных чисел, матрица собственных векторов и матрица коэффициентов.

Полученные коэффициенты одного из преобразований или исходные сигналы использовались для обучения однослойного перцептрона, построенного аналогично тому, что использовался в первом модуле. Логистическая регрессия обучалась в реализаций: 1 против всех для случая 1 против 2 (3 класса: «позитивные», «негативные» и прочие мимические выражения, 1 против 4 (5

классов – «радость», «грусть», «удивление», «злость» и «усталость»), 1 против 13 и 1 против 15 (сокращенный и полный наборы базовых мимических выражений).

Результаты экспериментальной проверки алгоритмов

В результате экспериментальной проверки алгоритмов (см. табл 1) на выборках, содержащих представителей разных классов мимических выражений, получен результат для однослойного перцептрона для наихудшего случая порядка 98,7%. Для случаев 1 против 13 и 1 против 15 были получены неудовлетворительные результаты, что требует увеличения объема обучающей выборки для данной реализации алгоритма классификации и изменения архитектуры и типа классификатора.

Таблица 1. Точность идентификации примеров для различных реализаций для 3 классов мимических выражений

Метод	Реализации и число отсчетов в сигнале				
	Сигнал целиком				1 отсчет
	432	1080	2160	4320	36
Перцептрон	64,8%				98%
DCT+KLT	64,8%				-
SVD+KLT	55%	58%	60%	62,5%	-
SVD	55%	58%	59%	61%	32%
KLT/FFT+KLT	52%	54%	57%	59,5%	-

Анализ на различных выборках и различном количестве классов показал, что уменьшение размера обучающей выборки, как путем отсечения части сигнала, или числа обучающих примеров приводит к существенному снижению точности распознавания отдельных классов. Так, отсечение половины сигнала или уменьшение объема обучающей выборки с 150 до 110 примеров (и обучающей с 37 до 17 примеров), приводит к снижению точности до 40%; разбиение сигналов на окна половинной длины с шагом 5 отсчетов дало результаты – 50,2%.

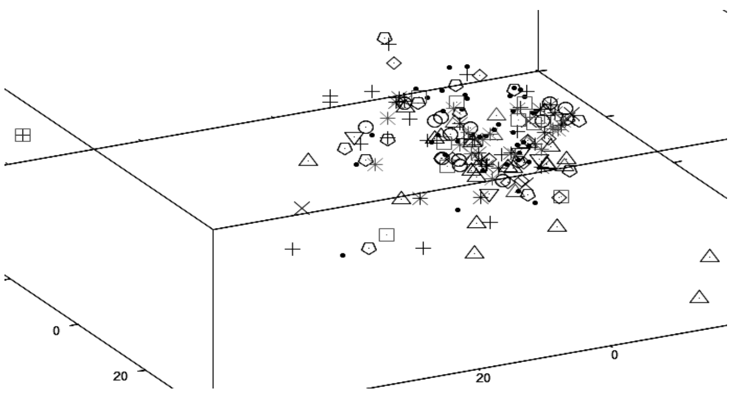


Рис. 5. Представление исходных сигналов в пространстве собственных векторов (№ 2,3 и 4) преобразования KLT для всего сигнала (14 классов).

Причиной низкого процента распознавания, скорее всего, является плохая разделимость классов в пространстве первых трех-четырех собственных векторов интегральных преобразований (рис. 5).

Выводы

В результате исследований была получена обучающая выборка, содержащая 187 примеров различных мимических выражений. Обучающая выборка использовалась в двух различных реализациях алгоритма классификации мимических выражений, использующую, в качестве обучающих примеров, состояния лица в отдельно взятый момент времени и для всего интервала мимического проявления для всех примеров. Результаты проверки алгоритмов классификации показали, что один отсчет данных может быть использован для априорной классификации мимических выражений на несколько базовых классов (от 2 до 16 классов) на основе алгоритма однослойного перцептрона. Для получения хороших результатов работы алгоритма обучения на основе всего сигнала с использованием интегральных преобразований необходимо увеличение объема обучающей выборки и применение многослойной нейросети или другого типа классификатора.

В дальнейших исследованиях предложено проверить алгоритмы классификации на отдельных микровыражениях и, на основе идентифицированных микровыражений, идентифицировать мимические выражения.

Литература

1. Зайцева Г.Л. Жестовая речь. Дактилология: Учеб. для студ. высш. учеб. заведений / Зайцева Г.Л. – М: Гуманит. изд. центр ВЛАДОС, 2000. – 192с.
2. Ткачева Т. Жестовый язык глухих как особая знаковая коммуникативная система (на материале русского и французского жестовых языков) / Ткачева Т., Ахмедышева Н. // Материалы Международной научно-практической конференции «Приоритеты и интересы современного общества». – Астрахань: АГУ, 2010. – С.361-365.
3. Кульбіда С.В. Українська жестова мова як природна знакова система / С. В. Кульбіда // збірник наукових праць «Жестова мова й сучасність»: К: Педагогічна думка, 2009 . – С. 218-239.
4. Крак Ю.В., Тернов А.С., Кузнєцов В.А. Разработка аудиовизуальной базы данных немануальных компонент жестовой речи // Управляющие системы и машины (УСиМ). – 2013. – №4. – С.58-62.
5. Крак Ю.В., Кузнецов В.А., Тернов А.С. Моделирование эмоциональных проявлений с помощью параметрических моделей головы человека // Штучний інтелект. – 2013. – №3(61). – С.154-170.
6. Dailey Matthew N. EMPATH: A Neural Network that Categorizes Facial Expressions / Matthew N. Dailey, Garrison W. Cottrell, Curtis Padgett, Ralph Adolphs// Massachusetts Institute of Technology Journal of Cognitive Neuroscience, – 2002. – Vol. 14, No 8. – P.1158-1173.
7. Krishna K.V., Varma K., Varma G. Hybrid Emotional Neural Network for Facial Expression Classification // International Journal of Computer Applications – Foundation of Computer Science, New York. – 2011. –Vol. 35, No 12. – P. 8-22.
8. Novakovic J. et al. Classification Accuracy of Neural Networks with PCA in Emotion Recognition / Jasmina Novakovica, Milomir Minica, Alempije Veljovicb // Theory and Applications of Mathematics & Computer Science – University of Arad, Arad, Romania. – 2011. – №1. – P.11-16.
9. M. Karthigayan et al. Genetic Algorithm and Neural Network for Face Emotion Recognition // M. Karthigayan, M. Rizon, R. Nagarajan and Sazali Yaacob / Affective Computing - I-Tech Education and Publishing. – 2008. – P. 57-67.
10. Леонтьев В.О. Классификация эмоций / Леонтьев В.О. – Одесса: Инновационно-ипотечный центр, 2002. - 84 с.

11. Изард К.Е. Эмоции человека / Изард К.Е. – М.: МГУ, 1980. – 439с.
12. Ekman P. Unmasking the face. A guide to recognizing emotions from facial clues / Cambridge, M.A., Malor books, 2003. – 214 p.
13. Дарвин Ч. Выражение эмоций у человека и у животных /Дарвин Ч. // Сочинения; [пер. С.Л.Соболя под ред. акад. Е.Н. Павловского: в 8 т.] – М.: Изд. АН СССР – 1953. – Т.5. – С.431-480.
14. Medlej J. Drawing People / Medlej J. - London.: Cedarseed, 2012. – 99 p.
15. Iurii Krak, Anton Ternov, Vladislav Kuznetzov. Facial expressions analysis based on a computer vision algorithms //International Journal Information Models and Analysis. –2014. – Vol.1.– P.62-77.
16. Alberto Menache. Understanding Motion Capture for Computer Animation and Video Games //Alberto M.// - Morgan Kaufmann Publishers Inc. San Francisco, CA, 1999. – 231 p.
17. Miller F. P. Facial Action Coding System. / Miller F. P., Vandome A. F., McBrewster J. – Beau Bassin, International Book Marketing Service Ltd, 2011. – 80 p.
18. Matthias Grundmann et al. Auto-Directed Video Stabilization with Robust L1 Optimal Camera Paths // Matthias Grundmann, Vivek Kwatra, Irfan Essa // IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2011) – 2011. - Colorado Springs, USA. – P. 225-232.
19. Grewal M.S., Andrews A.P. Kalman Filtering – Theory and Practice Using MATLAB. – Editions Wiley, 2001. – 640 p.
20. Электронный ресурс библиотеки компьютерного зрения OpenCV. Режим доступа: http://opencv.org/
21. Электронный ресурс библиотеки перекодировщика видео avconv/ffmpeg. Режим доступа: http://libav.org/news.html
22. Электронный ресурс приложения Blender. Режим доступа: http://www.blender.org
23. Нуссбаумер Г. Быстрое преобразование Фурье и алгоритмы вычисления сверток: Пер. с англ. - М.: Радио и связь, 1985. – 248 с.
24. Flannery Rian P. et al. Numerical Recipes Example Book // Rian P. Flannery, Saul A. Teukolsky, William H. Press, William T. Vetterling. – Cambridge University Press, 1992. – 336 p.
25. Chumakov A.G., Kurashov V.N. Karhunen-Loeve Basis Synthesis for Grate Capacity Signal Performance in Optical Processors // SPIE Proc. – 1993. – Vol.21-08. – P.338- 342.
26. Loeffler C., Ligtenberg A., Moschytz G.. Practical Fast 1-D DCT Algorithms with 11 Multiplications // Proc. Int. Conf. on Acoustics, Speech, and Signal Processing 1989 (ICASSP '89). – 1989. – P. 988-991.
27. Электронный ресурс программы Octave. Режим доступа: https://www.gnu.org/software/octave/

Literatura

1. Zaytseva G.L. Zhestovaya rech. Daktilologiya: Ucheb. dlya stud. vyissh. ucheb. zavedeniy / Zaytseva G.L. – M: Gumanit. izd. tsentr VLADOS, 2000. – 192s.
2. Tkacheva T. Zhestovyiy yazyik gluhih kak osobaya znakovaya kommunikativnaya sistema (na materiale russkogo i frantsuzskogo zhestovyih yazyikov) / Tkacheva T., Ahmedyisheva N. // Materialyi Mezhdunarodnoy nauchno-prakticheskoy konferentsii «Prioritetyi i interesyi sovremennogo obschestva». – Astrahan: AGU, 2010. – S.361-365.
3. Kulbida S.V. UkraYinska zhestova mova yak prirodna znakova sistema / S. V. KulbIda // zbIrnik naukovih prats «Zhestova mova y suchasnIst»: K: PedagogIchna dumka, 2009 . – S. 218-239.
4. Krak Yu.V., Ternov A.S., KuznEtsov V.A. Razrabotka audiovizualnoy bazyi dannyih nemanualnyih komponent zhestovoy rechi // Upravlyayuschie sistemyi i mashinyi (USiM). – 2013. – #4. – S.58-62.
5. Krak Yu.V., Kuznetsov V.A., Ternov A.S. Modelirovanie emotsionalnyih proyavleniy s pomoschyu parametricheskih modeley golovyi cheloveka // Shtuchniy Intelekt. – 2013. – #3(61). – S.154-170.
6. Dailey Matthew N. EMPATH: A Neural Network that Categorizes Facial Expressions / Matthew N. Dailey, Garrison W. Cottrell, Curtis Padgett, Ralph Adolphs// Massachusetts Institute of Technology Journal of Cognitive Neuroscience, – 2002. – Vol. 14, No 8. – R.1158-1173.
7. Krishna K.V., Varma K., Varma G. Hybrid Emotional Neural Network for Facial Expression Classification // International Journal of Computer Applications – Foundation of Computer Science, New York. – 2011. –Vol. 35, No 12. – R. 8-22.
8. Novakovic J. et al. Classification Accuracy of Neural Networks with PCA in Emotion Recognition / Jasmina Novakovica, Milomir Minica, Alempije Veljovicb // Theory and Applications of Mathematics & Computer Science – University of Arad, Arad, Romania. – 2011. – #1. – R.11-16.

9. M. Karthigayan et al. Genetic Algorithm and Neural Network for Face Emotion Recognition // M. Karthigayan, M. Rizon, R. Nagarajan and Sazali Yaacob / Affective Computing - I-Tech Education and Publishing. – 2008. – R. 57-67.
10. Leontev V.O. Klassifikatsiya emotsiy / Leontev V.O. – Odessa: Innovatsionno-ipotechnyii tsentr, 2002. 84 s.
11. Izard K.E. Emotsii cheloveka / Izard K.E. – M.: MGU, 1980. – 439s.
12. Ekman P. Unmasking the face. A guide to recognizing emotions from facial clues / Cambridge, M.A., Malor books, 2003. – 214 r.
13. Darvin Ch. Vyirazhenie emotsiiy u cheloveka i u zhivotnyih. / Darvin Ch. // Sochineniya; [per. S.L.Sobolya pod red. akad. E.N. Pavlovskogo: v 8 t.] – M.: Izd. AN SSSR – 1953. – T.5. – C.431-480.
14. Medlej J. Drawing People / Medlej J. - London.: Cedarseed, 2012. – 99 p.
15. Iurii Krak , Anton Ternov, Vladislav Kuznetzov. Facial expressions analysis based on a computer vision algorithms //International Journal Information Models and Analysis. –2014. – Vol.1.– P.62-77.
16. Alberto Menache. Understanding Motion Capture for Computer Animation and Video Games //Alberto M.// - Morgan Kaufmann Publishers Inc. San Francisco, CA, 1999. – 231 p.
17. Miller F. P. Facial Action Coding System. / Miller F. P., Vandome A. F., McBrewster J. – Beau Bassin, International Book Marketing Service Ltd, 2011. – 80 p.
18. Matthias Grundmann et al. Auto-Directed Video Stabilization with Robust L1 Optimal Camera Paths // Matthias Grundmann, Vivek Kwatra, Irfan Essa // IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2011) – 2011. - Colorado Springs, USA. – P. 225-232.
19. Grewal M.S., Andrews A.P. Kalman Filtering – Theory and Practice Using MATLAB. – Editions Wiley, 2001. – 640 p.
20. Elektronnyiy resurs biblioteki kompyuternogo zreniya OpenCV. Rezhim dostupa: http://opencv.org/
21. Elektronnyiy resurs biblioteki perekodirovschika video avconv/ffmpeg. Rezhim dostupa: http://libav.org/news.html
22. Elektronnyiy resurs prilozheniya Blender. Rezhim dostupa: http://www.blender.org
23. Nussbaumer G. Byistroe preobrazovanie Fure i algoritmyi vyichisleniya svertok: Per. s angl. - M.: Radio i svyaz, 1985. – 248 s.
24. Flannery Rian P. et al. Numerical Recipes Example Book // Rian P. Flannery, Saul A. Teukolsky, William H. Press, William T. Vetterling. – Cambridge University Press, 1992. – 336 p.
25. Chumakov A.G., Kurashov V.N. Karhunen-Loeve Basis Synthesis for Grate Capacity Signal Performance in Optical Processors // SPIE Proc. – 1993. – Vol.21-08. – P.338- 342.
26. Loeffler C., Ligtenberg A., Moschytz G.. Practical Fast 1-D DCT Algorithms with 11 Multiplications // Proc. Int. Conf. on Acoustics, Speech, and Signal Processing 1989 (ICASSP '89). – 1989. – P. 988-991.
27. Elektronnyiy resurs programmyi Octave. Rezhim dostupa: https://www.gnu.org/software/octave/

RESUME

A.S. Ternov, V.A. Kuznetsov, Ju.V. Krak

Analysis of facial expressions on a human face

Described in the article are the steps of data acquisition and creation of facial expressions analysis toolkit based on classification algorithms. In order to get the data, facial expressions of three actors were recorded on video. Then, videos which contained facial expressions were processed by computer vision algorithms. In order to improve effectiveness of computer vision algorithms in processing the video data, videos and motion data acquired from these algorithms were processed by two step noise reduction algorithm, included digital video stabilizer and Cullman filter. Motion data, that describe temporal characteristics of facial expressions, formed teach and control data for classification algorithms.

These algorithms included 2 different versions of algorithms. Both versions implemented contain 1-layer perceptron; the main difference is data – first version utilized every single frame of signal as single teach or control data element, the second utilized DCT, FFT, KLT and SVD decomposition coefficients of whole signal,

described change of all the facial features during 1 video fragment. Different teach and control data sets were used to test these algorithms; quantity of classes differs – from 2 to 16: a) 3 classes - (positive and negative facial expressions), b) 5 classes - «happiness», «sadness», «wondering», «anger» и «tiredness»; c) 14 classes; d) 16 classes.

Results on test datasets proof that one-layer perceptron can be easily applied for single frame facial expressions marker coordinates with quality approx. 98,7%, although decomposition coefficients (while utilizing DCT, FFT, KLT and SVD decomposition) of whole signal give mediocre quality – 64,8% on hybrid DCT-KLT algorithm, show better results require increase of dataset size and type of classifier.

This research is to be extended on facial microexpressions, what can help to classify entire facial expressions based on identified facial microexpressions.

Ю.В. Крак, В.А. Кузнєцов, А.С. Тернов

Аналіз мімічних проявів на обличчі людини

У статті описано етапи отримання даних і створення інструментів аналізу мімічних проявів за допомогою алгоритмів класифікації. Для цього було записано відеофрагменти мімічних проявів для трьох різних акторів. Відеофрагменти, які містили мімічні прояви, оброблялися алгоритмами комп'ютерного зору. З метою покращення роботи алгоритмів комп'ютерного зору було застосовано двоступінчастий алгоритм зниження шумів у даних, що включав в себе алгоритм цифрової стабілізації відео і фільтр Калмана.

Отримані числові дані, що містили координати маркерів на обличчі актора під час мімічного прояву для всіх відеофрагментів, формували навчальну і контрольну вибірки. Було створено 2 реалізації алгоритму класифікації мімічних виразів. Перша містила одношаровий перцептрон, навчений для чотирьох випадків: а) 3 класи «позитивні» і «негативні» мімічні вирази; б) 5 класів – «радість», «смуток», «подив», «злість» і «втома»; в) 14 класів; г) 16 класів. Для навчання вибиралися стани маркерів в окремо взятий проміжок часу. Друга містила одношаровий перцептрон, що навчався на коефіцієнтах інтегральних перетворень: Карунена-Лоева, косинусного, сингулярного розкладу і Фур'є. Коефіцієнти інтегральних перетворень розраховувались на основі станів маркерів на всьому інтервалі мімічного прояву.

У результаті експериментальної перевірки алгоритмів класифікації на вибірках, що містили представників різних класів мімічних виразів, отримано результат порядку 98,7% з використанням класифікатора на базі перцептрона, що вказує на можливість класифікації мімічних виразів на основі одиничного відліку. При використанні гібридного алгоритму DCT-KLT і класифікатора на базі перцептрона точність склала 64,8%, що вказує на необхідність розширення навчальної вибірки і вибору іншого типу класифікаторів.

Запропоновано перевірити алгоритми класифікації на окремих мікровиразах і, на основі ідентифікованих мікровиразів ідентифікувати мімічні вирази.

Поступила в редакцию 11.09.2015

